Talk at CENDI/NFAIS workshop Open Science: Driving Forces and Practical Realities, Washington, DC, November 12, 2013

Open Research: Key Concepts and Issues

John Leslie King

jlking@umich.edu

University of Michigan

Leading change is difficult, perilous, and uncertain of success.

The innovator's enemies benefit from old conditions, have laws on their side, and attack like partisans.

The innovator's defenders are lukewarm, afraid of the enemies, and unaccepting of new things without experience of their benefits.

Nicollo Machiavelli, 500 years ago this year

Six Points

- Mandates will not work
- Compelling but not new
- Changes
- Interests served
- Culture
- What to do?

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20502

February 22, 2013

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

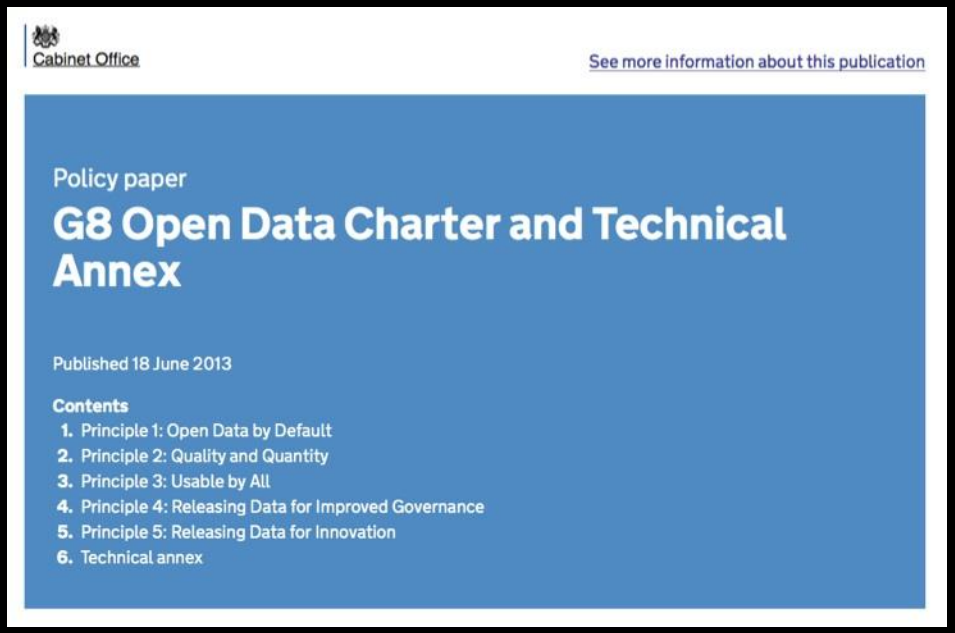
FROM: John P. Holdren
Director

SUBJECT: Increasing Access to the Results of Federally Funded Scientific Research

1. Policy Principles

The Administration is committed to ensuring that, to the greatest extent and with the fewest constraints possible and consistent with law and the objectives set out below, the direct results of federally funded scientific research are made available to and useful for the public, industry, and the scientific community. Such results include peer-reviewed publications and digital data.

Mandates will not work

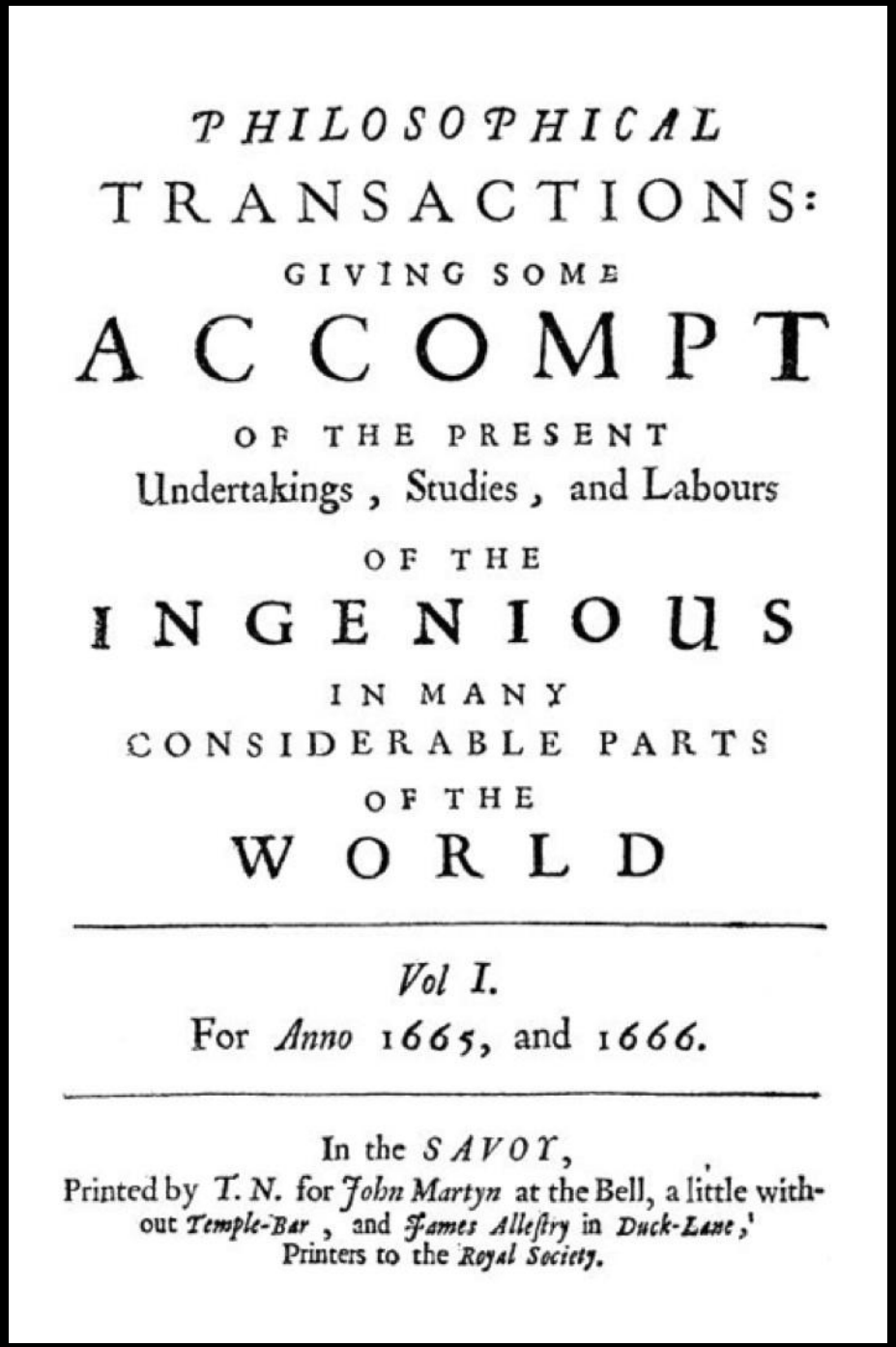


It's all over the world: the G8, Australia, the EU – this is just a few.

Mandates will not work

The video here was at this URL as of November 12, 2013:
https://www.youtube.com/watch?v=dkYfmRwryQo

Mandates will not work



PHILOSOPHICAL
TRANSACTIONS:
GIVING SOME
ACCOMPT
OF THE PRESENT
Undertakings, Studies, and Labours
OF THE
INGENIOUS
IN MANY
CONSIDERABLE PARTS
OF THE
WORLD

Vol I.
For Anno 1665, and 1666.

In the SAVOY,
Printed by T. N. for John Martyn at the Bell, a little without Temple-Bar, and James Allestry in Duck-Lane,
Printers to the Royal Society.

The Royal Society of London for Improving Natural Knowledge, an 'invisible college' of natural philosophers. First met in the mid-1640s, official founding in 1660. Notable names: Christopher Wren, Robert Hooke, Robert Boyle, and Isaac Newton. Henry Oldenburg edited the first issue of *Philosophical Transactions.*

Compelling but not new

Through “...a vast commerce of ideas…” knowledge “…triumphed over distance, over differences of language, over diversity of habits, over prejudice, and over bigotry.”

Daniel Webster, June 17, 1825
Groundbreaking for the Bunker Hill Monument

The pursuit of knowledge: from pastime of the curious into a dynamo of social and economic advancement.

Compelling but not new

equipment, and to Dr. C. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

[1] Young, F. B., Gerrard, H., and Jevons, W., *Phil. Mag.*, **40**, 149 (1920).

[2] Longuet-Higgins, M. S., *Mon. Not. Roy. Astro. Soc., Geophys. Supp.*, **5**, 285 (1949).

[3] Von Arx, W. S., Woods Hole Papers in Phys. Oceanog. Meteor., **11** (1) (1950).

[4] Ehrman, V. W., *Arkiv. Mat. Astron. Fysik. (Stockholm)*, **2** (11) (1905).

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

This figure is purely diagrammatic. The two ribbons symbolize the two phosphate–sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3′,5′ linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4 A. in the *z*-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical *z*-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric form (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at King's College, London. One of us (J. D. W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J. D. WATSON
F. H. C. CRICK
Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems, Cavendish Laboratory, Cambridge.
April 2.

[1] Pauling, L., and Corey, R. B., *Nature*, **171**, 346 (1953); *Proc. U.S. Nat. Acad. Sci.*, **39**, 84 (1953).

[2] Furberg, S., *Acta Chem. Scand.*, **6**, 634 (1952).

[3] Chargaff, E., for references see Zamenhof, S., Brawerman, G., and Chargaff, E., *Biochim. et Biophys. Acta*, **9**, 402 (1952).

[4] Wyatt, G. R., *J. Gen. Physiol.*, **36**, 201 (1952).

[5] Astbury, W. T., Symp. Soc. Exp. Biol. 1, Nucleic Acid, 66 (Camb. Univ. Press, 1947).

[6] Wilkins, M. H. F., and Randall, J. T., *Biochim. et Biophys. Acta*, **10**, 192 (1953).

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

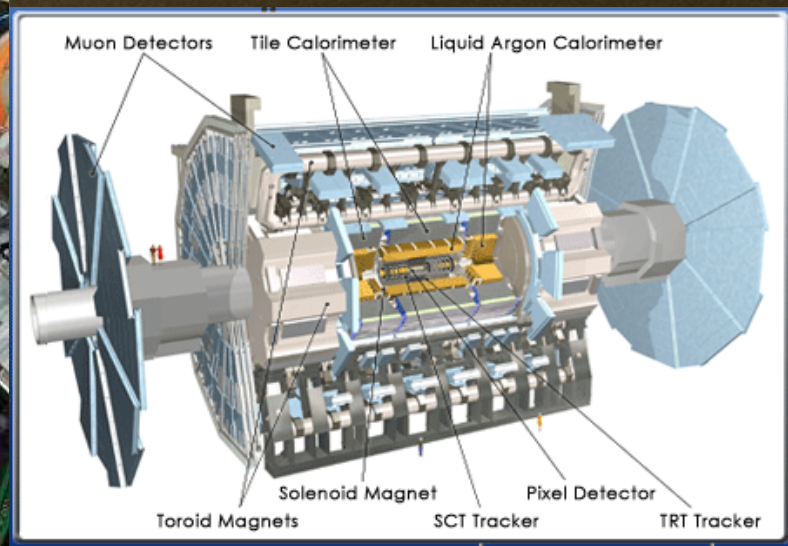
A Structure for Deoxyribose Nucleic Acid

J. D. WATSON
F. H. C. CRICK
Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems, Cavendish Laboratory, Cambridge.
April 2.

April 25, 1953, less than 2000 words

Changes


Science
21 December 2012 | $10
BREAKTHROUGH
of the YEAR
The HIGGS
BOSON
Changes
AAAS


Muon Detectors
Tile Calorimeter
Liquid Argon Calorimeter
Solenoid Magnet
Pixel Detector
Toroid Magnets
SCT Tracker
TRT Tracker

The Higgs Boson

The standard model of particle physics describes how elementary particles and a set of forces between them lead to all matter and most higher interactions, thereby providing a basis for understanding much of physics and chemistry. A key prediction of the standard model—that the universe is pervaded by a field that conveys mass—was tested and confirmed this year with the discovery of a particle associated with that field: the Higgs boson. Two separate, complex detectors housed in the largest and most energetic particle accelerator, the Large Hadron Collider (LHC) at CERN near Geneva, Switzerland, identified the characteristic decay products of the Higgs, allowing reconstruction of its mass. These experiments and this discovery were made possible by decades of cutting-edge,

ARTICLE

Journey in the Search for the Higgs Boson: The ATLAS and CMS Experiments at the Large Hadron Collider

M. Della Negra,[1] P. Jenni,[2] T. S. Virdee[1]*

The search for the standard model Higgs boson at the Large Hadron Collider (LHC) started more than two decades ago. Much innovation was required and diverse challenges had to be overcome during the conception and construction of the LHC and its experiments. The ATLAS and CMS Collaboration experiments at the LHC have discovered a heavy boson that could complete the standard model of particle physics.

[1]Imperial College, London SW7 2AZ, UK, and CMS, Geneva, Switzerland. [2]CERN, Geneva, Switzerland, and ATLAS, Geneva, Switzerland.

*To whom correspondence should be addressed. E-mail: tejinder.virdee@cern.ch

A New Boson with a Mass of 125 GeV Observed with the CMS Experiment at the Large Hadron Collider

The CMS Collaboration*†

The Higgs boson was postulated nearly five decades ago within the framework of the standard model of particle physics and has been the subject of numerous searches at accelerators around the world. Its discovery would verify the existence of a complex scalar field thought to give mass to three of the carriers of the electroweak force—the W^+, W^-, and Z^0 bosons—as well as to the fundamental quarks and leptons. The CMS Collaboration has observed, with a statistical significance of five standard deviations, a new particle produced in proton-proton collisions at the Large Hadron Collider at CERN. The evidence is strongest in the diphoton and four-lepton (electrons and/or muons) final states, which provide the best mass resolution in the CMS detector. The probability of the observed signal being due to a random fluctuation of the background is about 1 in 3×10^6. The new particle is a boson with spin not equal to 1 and has a mass of about 1.25 giga–electron volts. Although its measured properties are, within the uncertainties of the present data, consistent with those expected of the Higgs boson, more data are needed to elucidate the precise nature of the new particle.

*To whom correspondence should be addressed. E-mail: cms-spokesperson@cern.ch
†The complete list of authors and affiliations is available as supplementary material on *Science* Online.

A Particle Consistent with the Higgs Boson Observed with the ATLAS Detector at the Large Hadron Collider

The ATLAS Collaboration*†

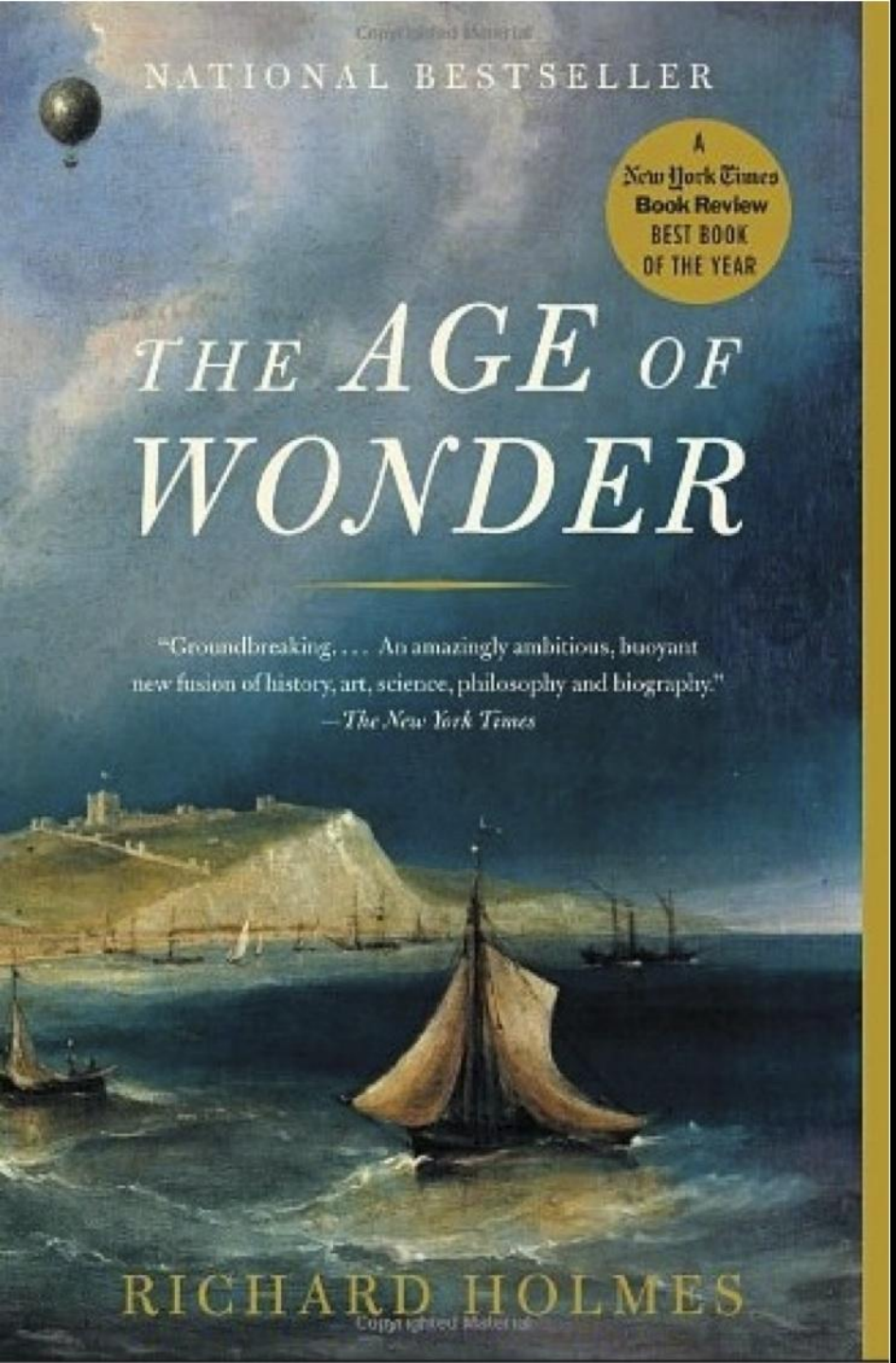
Nearly 50 years ago, theoretical physicists proposed that a field permeates the universe and gives energy to the vacuum. This field was required to explain why some, but not all, fundamental particles have mass. Numerous precision measurements during recent decades have provided indirect support for the existence of this field, but one crucial prediction of this theory has remained unconfirmed despite 30 years of experimental searches: the existence of a massive particle, the standard model Higgs boson. The ATLAS experiment at the Large Hadron Collider at CERN has now observed the production of a new particle with a mass of 126 giga–electron volts and decay signatures consistent with those expected for the Higgs particle. This result is strong support for the standard model of particle physics, including the presence of this vacuum field. The existence and properties of the newly

*To whom correspondence should be addressed. E-mail: atlas.publications@cern.ch
†The complete author list is included as supplementary material on *Science* Online.

Changes

Changes

Changes



Richard Holmes

Changes

The Shelby Amendment
Public Law 105-277, 1998
Public can use Freedom of Information Act to obtain research data resulting from federally sponsored projects. The law is binding on institutions as "owners" of the data.

U.S. Chamber of Commerce applauded the legislation for providing access to data used to impose regulatory burdens:

The public deserves better assurance that data used to support regulations is reliable and is not manipulated to support a particular view.

Interests served

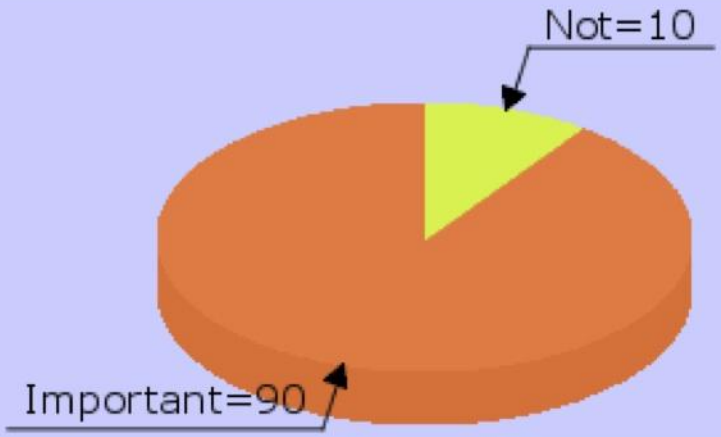
We have no idea what this will cost.

We have no idea about the benefits.

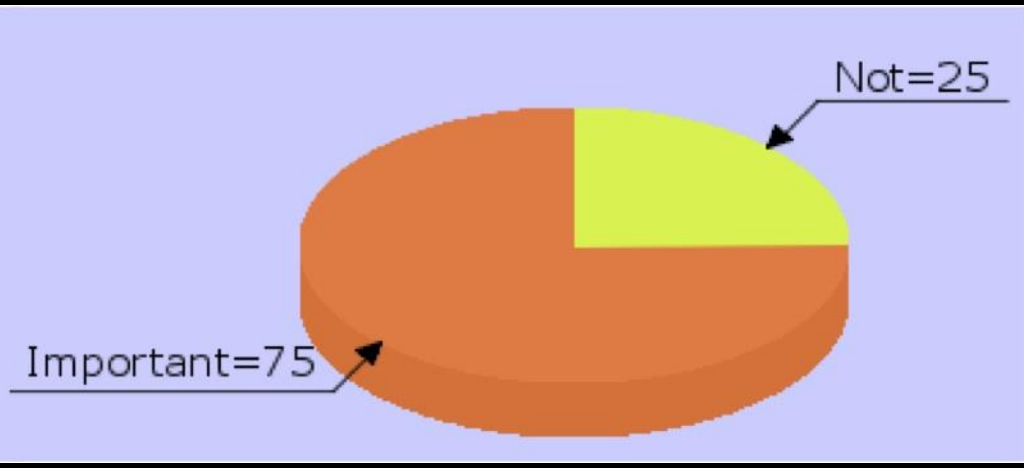
No one wants the costs on *them.*

Interests served

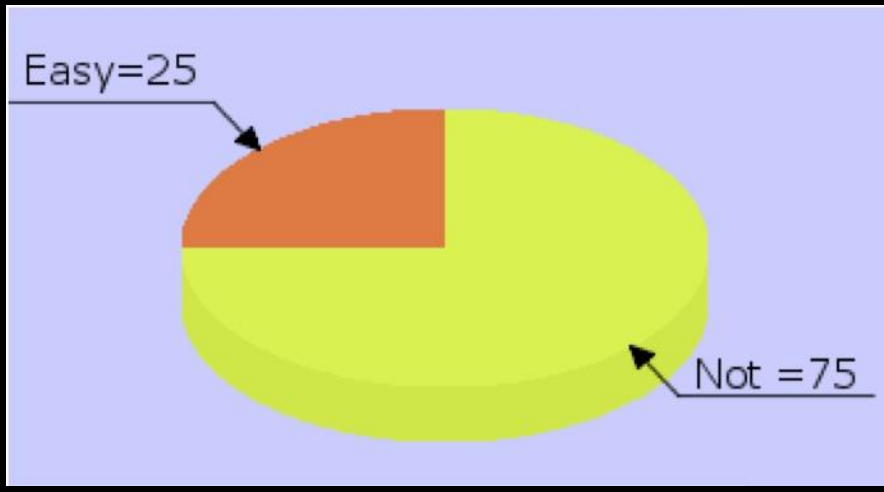
Open Sharing Important within area



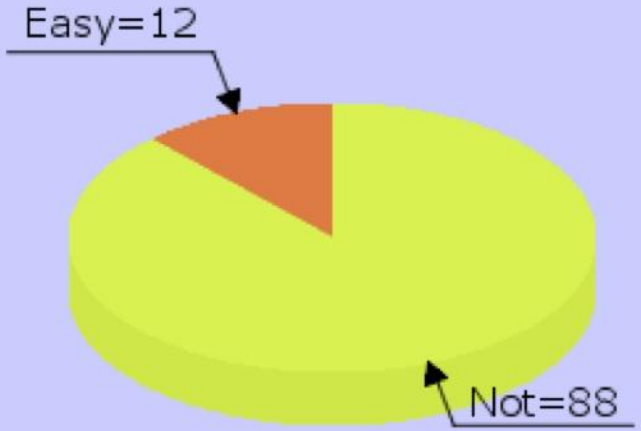
Open Sharing Important across areas



Open Sharing Easy within area



Open Sharing Easy across areas



Culture change

Burcu Bolukbasi, University of Illinois at Urbana-Champaign
Nicholas Berente, University of Georgia
Joel Cutcher-Gershenfeld, University of Illinois at Urbana-Champaign
Courtney G. Flint, Utah State University
Michael Haberman, University of Illinois at Urbana-Champaign
John Leslie King, University of Michigan
Eric Knight, University of Sydney
Barbara Lawrence, University of California at Los Angeles
Ethan Masella, Brandeis University
Charles Mcelroy, Case Western Reserve University, and John Carroll University
Barbara Mittleman, Nodality, Inc.
Mark Nolan, University of Illinois at Urbana-Champaign
Melanie Radik, Brandeis University
Namchul Shin, Pace University
Cheryl A. Thompson, University of Illinois at Urbana-Champaign
Susan Winter, University of Maryland
Ilya Zaslavsky, San Diego Super Computing Center, University of California at San Diego

Culture eats strategy for breakfast.

What To Do?

- Documentation
- Life cycle
- Workflow
- Metadata
- APIs
- Repositories
- Open Source
- Best Practices
- The Cloud
- Standards

Discussion.